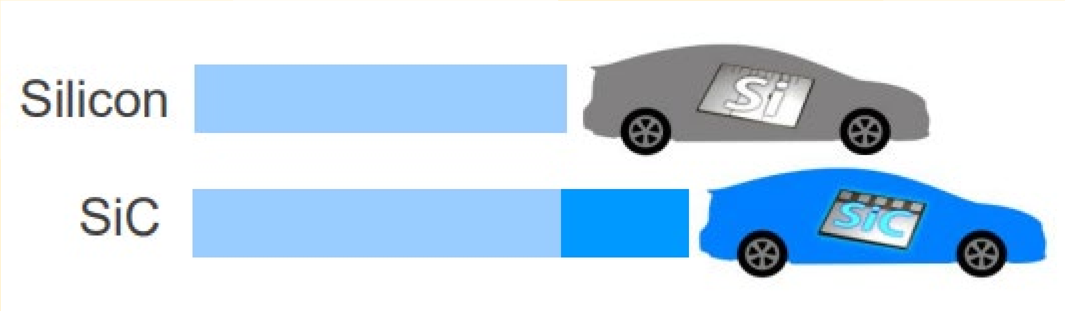
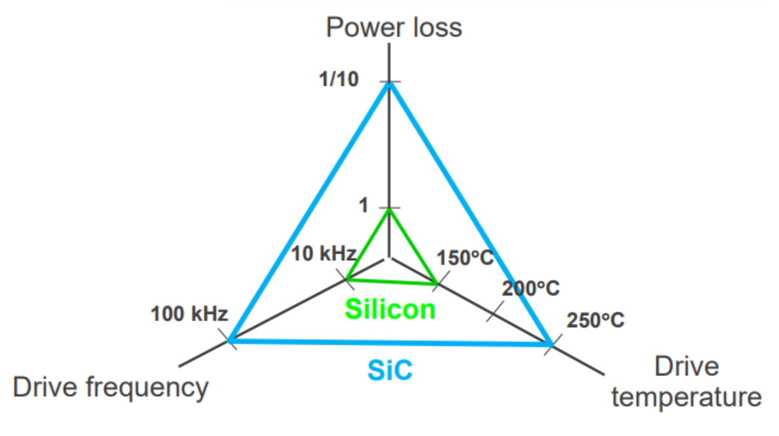
Bayesian Learning for Optimization and Analysis of SiC-based Inverter Package

Students: Hakki M. Torun, Ryan Wong
Faculty: Prof. Madhavan Swaminathan, Dr. Vanessa Smet
Collaborators: Toyota: Shailesh Joshi, Dr. Paul Fanson, Dr. Danny Lohan, Leo Liu

Objective



Metrics	Desired Metrics for Power Module
Size Reduction	> 25%
Power Density	100 kW per Wheel
Temperature	> 150°C
Carrier Frequency	> 10 kHz

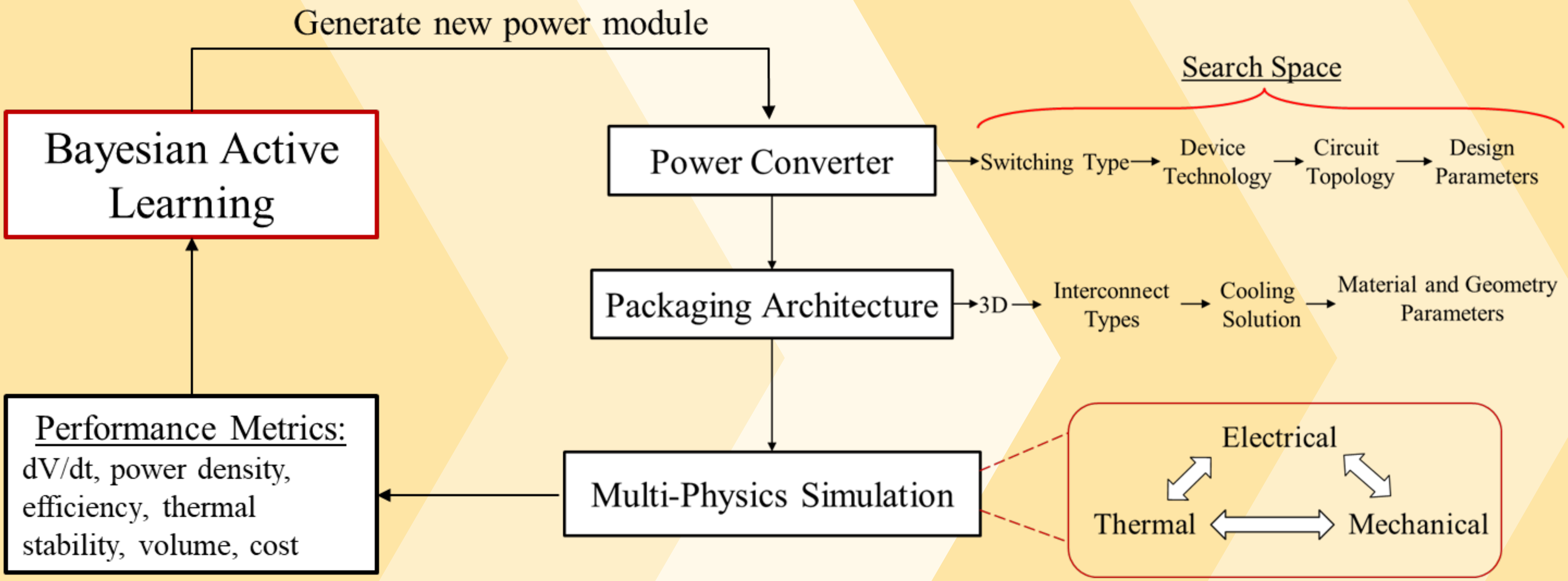


- Power module design complexity creates a gap between achieved and realizable performance.
  - Trade-offs regarding package architecture, circuit topology, materials and geometry.
- With various technologies available, how to determine which one is the best?
  - Can a new architecture be generated using the available ones?
- Objective is to use ML to determine optimal combination of package architecture, circuit topology and materials that will achieve the performance metrics for power module.

Acknowledgement

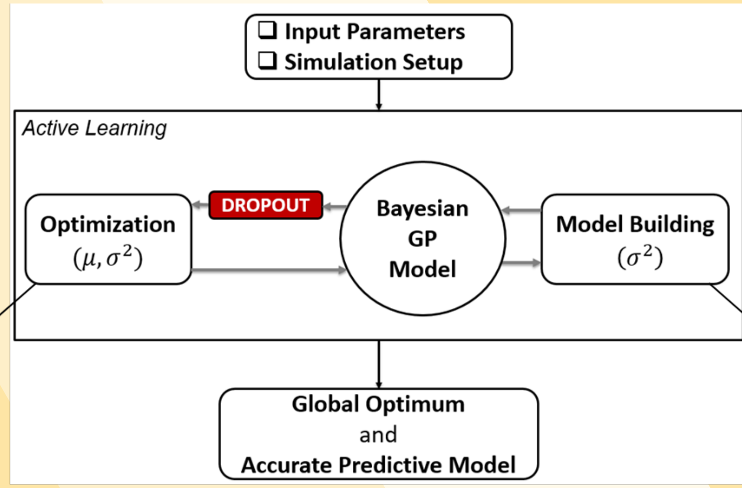
This project is funded by Toyota Motor Corporation.

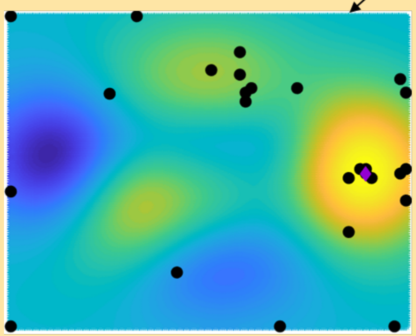
Machine Learning based DC/AC Power Module Optimization



- Power module optimization framework requires Design, Technology and Package Co-Optimization.
  - The best converter topology and package architecture combination, along with best design parameters.
- System is broken down to smallest possible building block at both circuit & package level.
- Key: Use Bayesian Active Learning (BAL) to determine the optimal combination of building blocks.
  - Along with quantifying the effect of choices on various performance metrics.

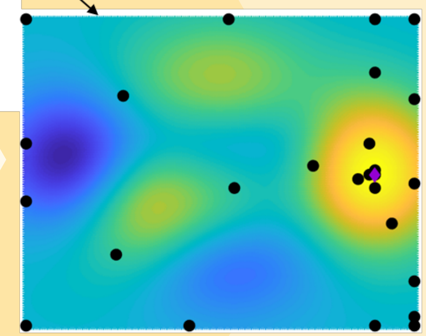
Bayesian Active Learning using Dropout (BALDO)





Design space explored for only for optimization

H. M. Torun et al, EPEPS'18

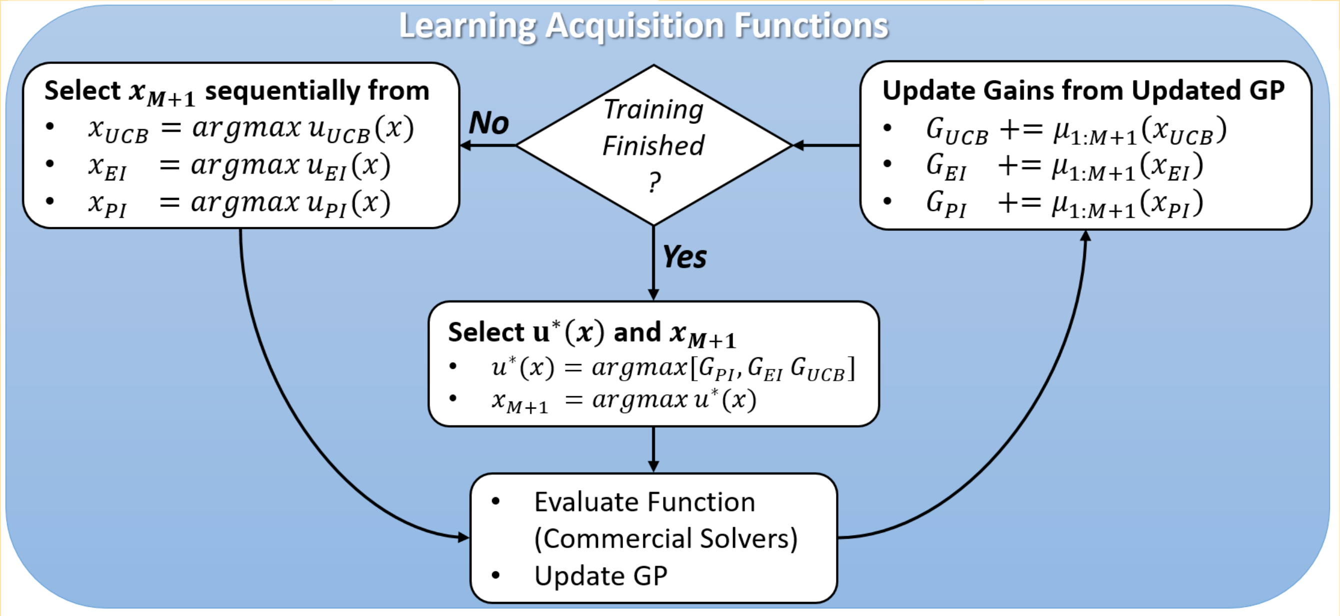


Design space explored for both optimization and model building

- ❑ Bayesian Optimization can find minimum temperature, but can't build an accurate model.
  - ❑ Need the model for sensitivity analysis.
- ❑ Is there a way to optimize AND build accurate model at the same time?
- ❑ Alternate between <u>optimization</u> and <u>model building</u> at every iteration.
  - ❑ **Better Model** → Avoid Local Optima → Faster Optimization
  - ❑ **Optimization** → Learn Saddle Points → Higher Model Quality

Complementary Objectives

Bayesian Active Learning using Dropout: Optimization Stage



Conventional BO:

❑Only use 1 acquisition function.

❑There is no guarantee that single acquisition function will outperform others at every problem.

Learning Acquisition Functions:

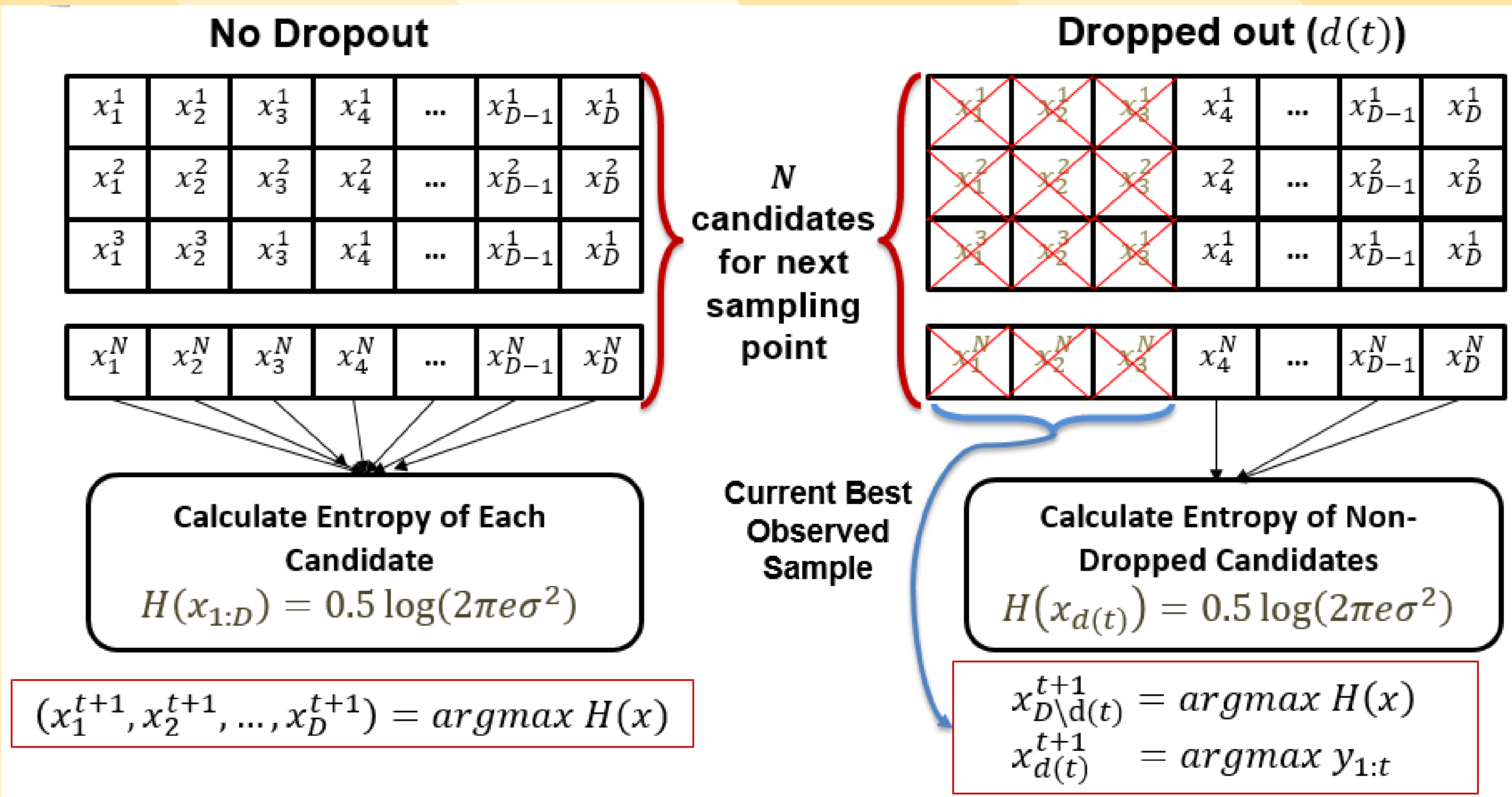
❑During optimization, actively learn which strategy is best for current problem.

❑After learning is completed, continue using the u(x) with highest gain.

❑Gains are updated even after learning is completed, hence, $u^*(x)$ is continuously updated.

[1]: H. M. Torun, M. Swaminathan, A. K. Davis, M. L. F. Bellaredj
"A Global Bayesian Optimization Algorithm and its Application to Integrated System Design". IEEE TVLSI'18.

Bayesian Active Learning using Dropout: Learning Stage



- The goal in learning stage is to select the sample such that decrement in uncertainty is maximized.
- We introduce dropout to prioritize optimization over learning.
- Only a group of dimensions (selected randomly at each iteration) are used for learning after dropout.
- Remaining dimensions are copied from the best observed sample so far.

H. Torun et al, EPEPS '18

Bayesian Training of Gaussian Process for Model Building

Two Approaches

Training for Optimization

Faster, approximate learning

Training for Model Building

Slower, accurate learning

Maximize Log Marginal Likelihood function

1. Non-Bayesian.
2. Gradient based optimization.
3. Non-convex problem (random restart).

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^n}{\operatorname{argmax}}\, p(D_{1:t} | y, m, \theta)$$

$$p\left(y^* \middle| x^*, D_{1:t}, m, \hat{\boldsymbol{\theta}}\right)$$

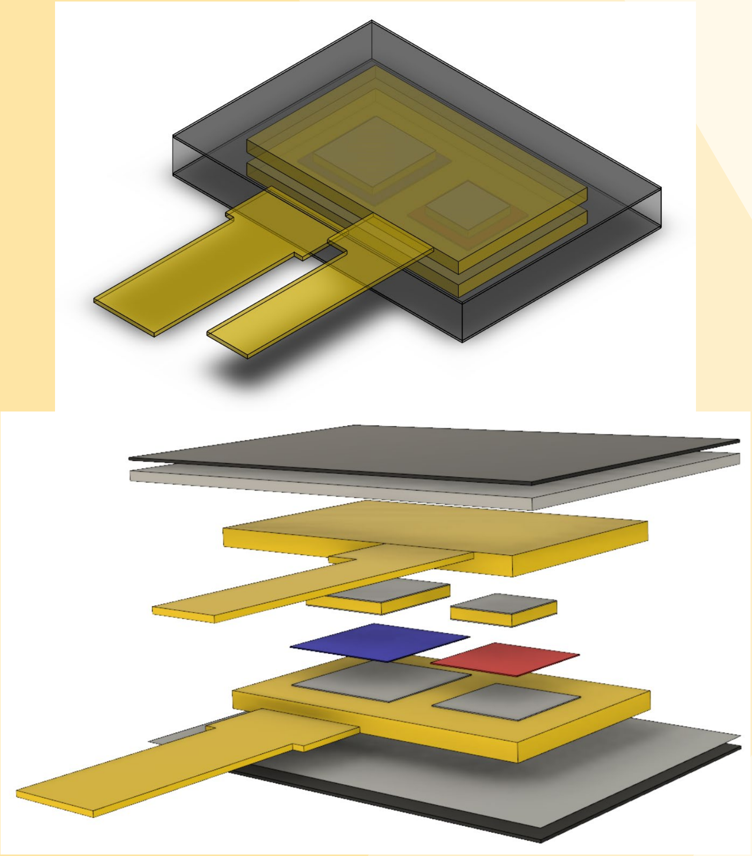
Derive Hyperparameter Posterior

1. Fully Bayesian.
2. Markov Chain Monte Carlo (MCMC) integration.
3. Integrate out all the uncertainties (ensemble).

Predictive Posterior

$$p(y^* | x^*, D_{1:t}, m) = \int p(y^* | x^*, D_{1:t}, m, \theta)\, \boldsymbol{p(\theta | D_{1:t}, m) d\theta}$$

Initial Problem: Thermal performance for DENSO Package



Parameter	Material	Unit	Min	Max
Diode/Switch Spacer Thickness	Cu	mm	0.20	3.00
Collector Plate Thickness	Cu	mm	0.05	3.00
Emitter Plate Thickness	Cu	mm	0.05	3.00
Collector Insulator Thickness	Silicon Nitride	mm	0.25	1.00
Emitter Insulator Thickness	Silicon Nitride	mm	0.25	1.00
All Joint Thicknesses (5 separate params.)	Solder	mm	0.05	0.10

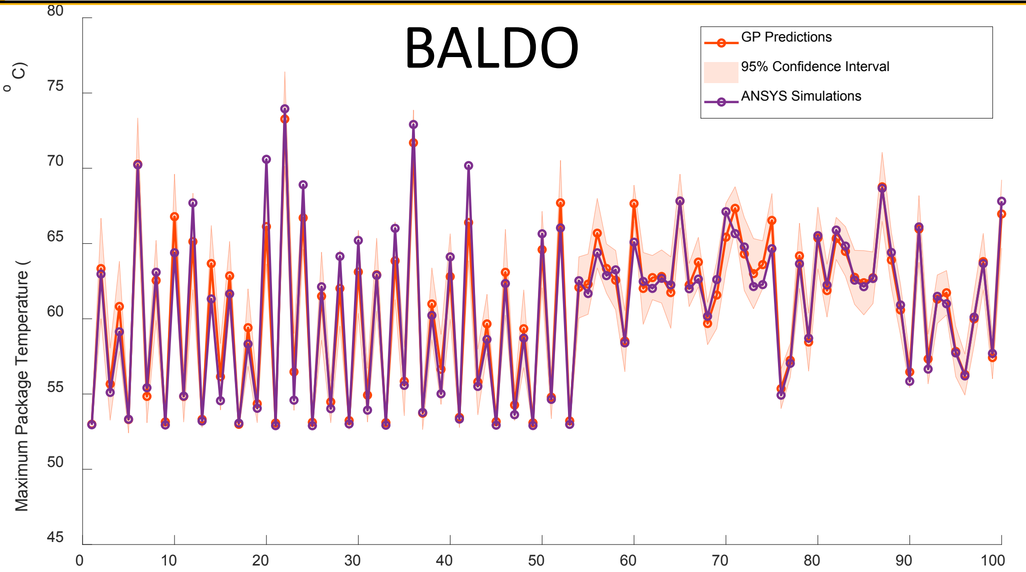
- ❑ 10 parameters define the geometry of the half-bridge rectifier.

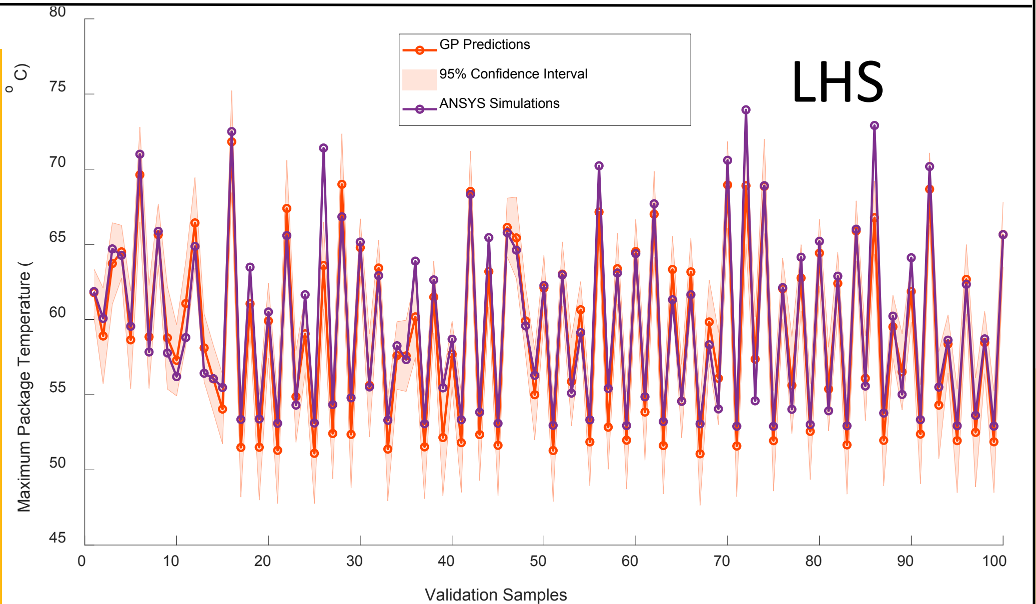
Objective:

1. Find thickness that minimizes maximum package temperature
2. Perform sensitivity analysis to determine which parameter(s) has more effect.

in minimum CPU time

Results: Model Accuracy



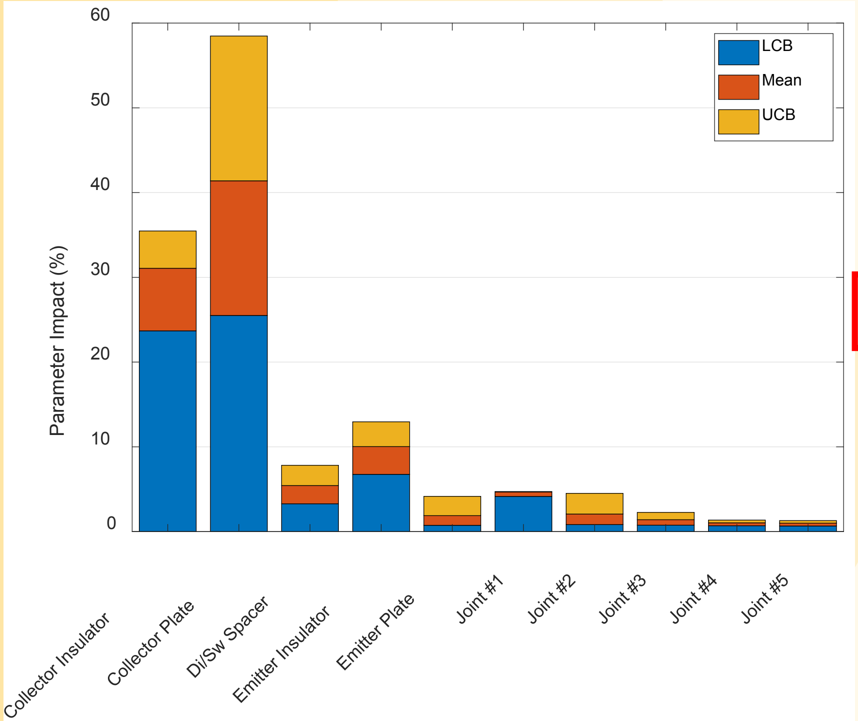


	LHS	BALDO
Norm. Mean Squared Err.	7.76%	4.94%
Max. Package Temperature	54.01°C	52.94°C
Av. Absolute Error	0.897°C	0.687°C
Max. Width of Confidence Interval	1.976°C	1.703°C

- Performance of models using data collected by BALDO and Latin Hypercube sampling is collected.
- For both methods, 50 samples are used for training and 100 for testing.

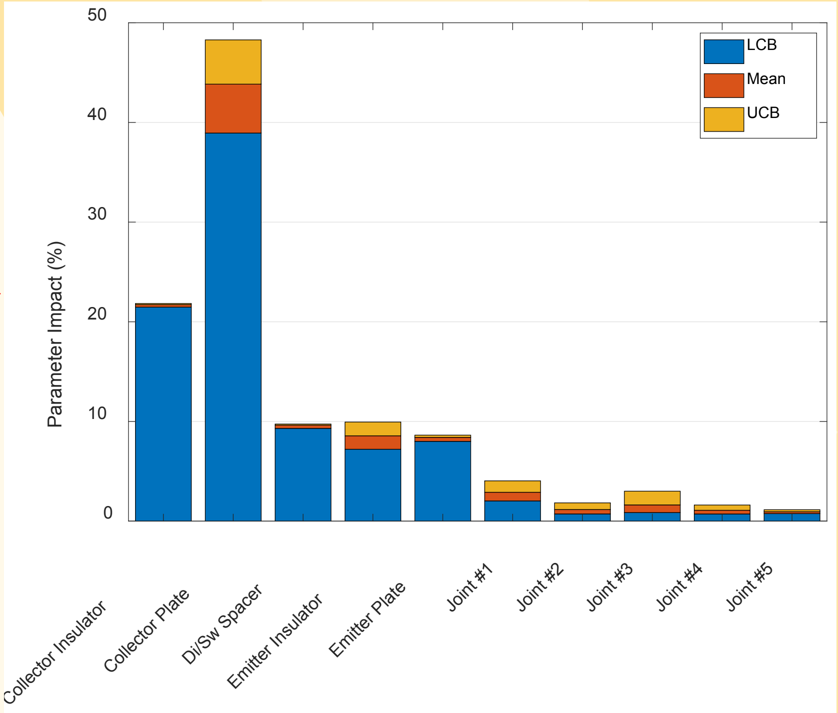
Results: Sensitivity Analysis

50 Samples by BALDO



MORE DATA

100 Samples by BALDO



- Sensitivity analysis is obtained as a by-product (free!) of active learning.
- As the data is scarce, confidence bounds over parameter weights are necessary.
  - Bayesian training of the GP allows to do so.
- As more data is added, confidence bounds shrink.
- Collector Plate thickness and Collector Insulator thickness has the largest impact.

Next Steps for ML

Circuit

Data Type: Continuous and categorical variables
Hard switching, soft switching, passives, device technology
Currently fixed as hard switching with IGBT devices (STEP IV)

Package Architecture

Data Type: Categorical variables
Wirebond, power card, 3D stacked, … + functional integration

NEXT STEP

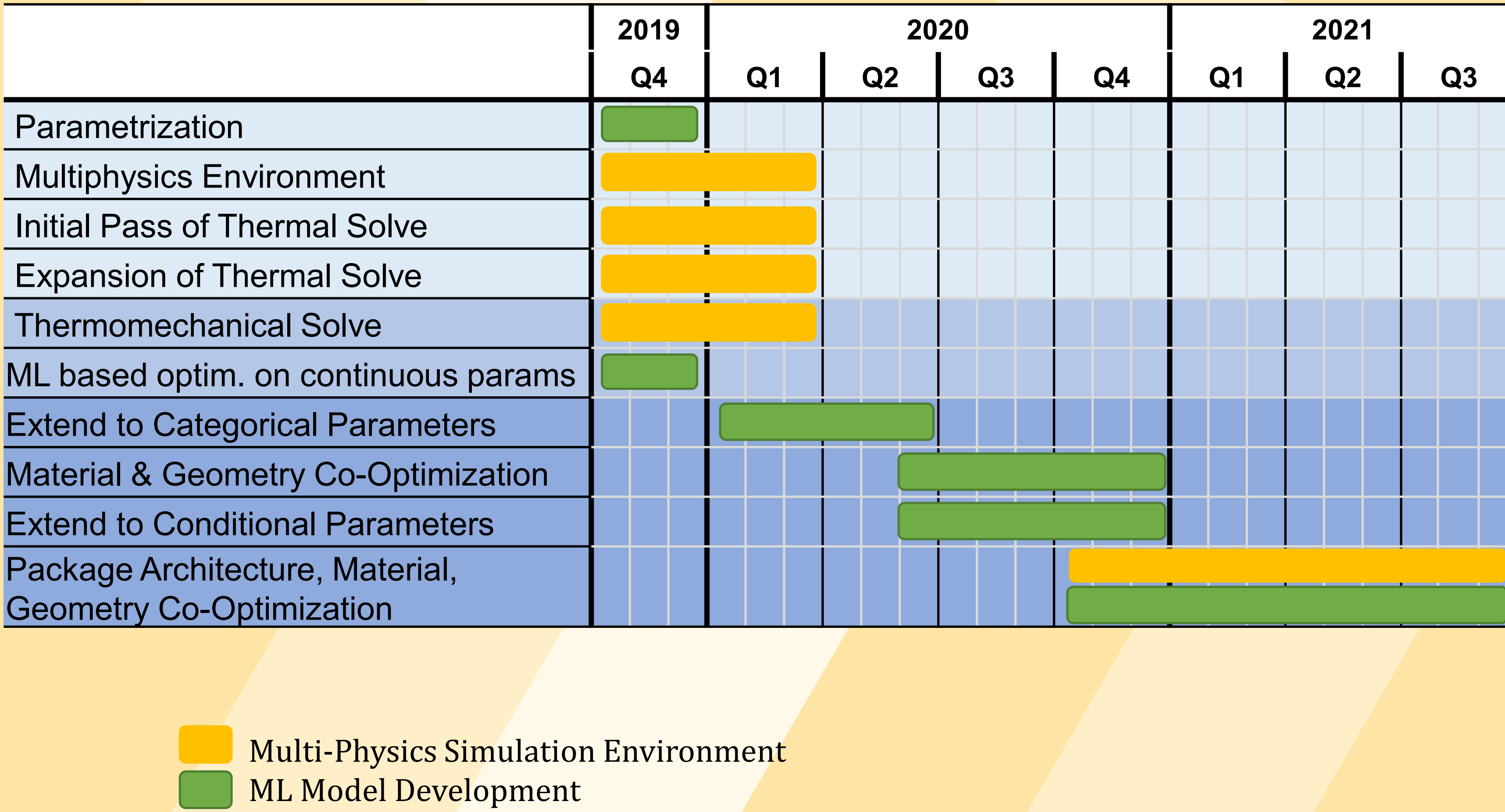
Geometry & Material

Data Type: Continuous and discrete variables
Current focus (STEP I): simplified problem to establish method capability
→ Denso package, single physics (thermal), 10 continuous variables (geometry: layer thicknesses)

Cooling

Data Type: Categorical and continuous
Heat pipe, heat sink, cold plate, integrated cooling, …
(Step III – explored in parallel through other Consortium project)

Timeline


2019
2020
2021
Q4
Q1
Q2
Q3
Q4
Q1
Q2
Q3
Parametrization
Multiphysics Environment
Initial Pass of Thermal Solve
Expansion of Thermal Solve
Thermomechanical Solve
ML based optim. on continuous params
Extend to Categorical Parameters
Material & Geometry Co-Optimization
Extend to Conditional Parameters
Package Architecture, Material, Geometry Co-Optimization
Multi-Physics Simulation Environment
ML Model Development

Summary

- Through this machine learning approach (BALDO), the power module structure (DENSO) was optimized for improved thermal performance (~2^{o}C).
- It can accommodate for up to 40 continuous parameters, with room for growth.
- It minimizes computational time and exhibits less error than other approaches.
- Critical elements in the module structure were highlighted for the most thermal impact (collector plate)